

Хорьков Д. А., аспирант
Гайдамакин Н. А., проф., д-р техн. наук

РЕЗУЛЬТАТЫ АНАЛИЗА ДЛИТЕЛЬНОСТЕЙ МЕЖСИМВОЛЬНЫХ ИНТЕРВАЛОВ В ЗАДАЧЕ АУТЕНТИФИКАЦИИ ПОЛЬЗОВАТЕЛЯ ПЭВМ ПО КЛАВИАТУРНОМУ ПОЧЕРКУ

В ходе первого года обучения в аспирантуре была проведена следующая работа. Выполнен библиографический поиск по доступным литературным источникам. В результате выяснено, что в открытой отечественной научной литературе материалов серьезных исследований феномена клавиатурного почерка не публиковалось. Большинство статей на русском языке представляют собой лишь обзор биометрических методов аутентификации и, в лучшем случае, содержат поверхностное упоминание математических методов, которые могут быть использованы при реализации систем аутентификации по клавиатурному почерку. В то же время в иностранной печати вопросы аутентификации по клавиатурному почерку освещены достаточно хорошо. Выяснено, что за рубежом разработки систем аутентификации по клавиатурному почерку ведутся с середины 1980-х годов. Несмотря на это, широкого внедрения коммерческих программных продуктов, реализующих аутентификацию по клавиатурному почерку, не происходит. Вместе с тем, возможность аутентификации по клавиатурному почерку, а в особенности — аутентификация по фрагменту произвольного текста, признается всеми авторами перспективной, но недостаточно изученной. Прежде всего это связано с окончательным переходом на подготовку всех видов документов при помощи ПЭВМ. Исходя из этого был сделан вывод о целесообразности проведения исследований по тематике аутентификации пользователя ПЭВМ по клавиатурному почерку в части непрерывной аутентификации по произвольному тексту.

Разработано программное обеспечение, позволяющее измерять длительность интервалов между нажатиями клавиш, а также продолжительность их удержания. В ходе обработки результатов измерений выяснено, что точность указанных измерений существенным образом зависит от модели используемой клавиатуры. Это обусловлено аппаратно-программным различием реализаций опроса клавиш микроконтроллером, встроенным в клавиатуру. Ни в одном из доступных литературных источников не было найдено ссылок на существование подобного ограничения, что говорит о недостаточной проработанности прикладных аспектов аутентификации по клавиатурному почерку.

С использованием указанного программного обеспечения были собраны и обработаны данные о длительностях межсимвольных интервалов и продолжительностях удержания клавиш несколькими пользователями при наборе различных текстов. Построены гистограммы распределений длительностей межсимвольных интервалов для наиболее часто встречающихся в русском языке пар символов (всего около 30 пар). На основании анализа этого

статистического материала сделан вывод о том, что распределения нельзя считать нормальными, поскольку они имеют явно асимметричную огибающую. Была сделана попытка аппроксимации этих распределений, в результате чего выяснено, что по своему виду они наиболее близки к логарифмически-нормальным. При качественном рассмотрении межсимвольных интервалов для конкретных пар клавиш в виде временного ряда было замечено, что в значениях наблюдаются «выбросы», которые имеет смысл считать случайными и исключить из рассмотрения. Эти выбросы являются, скорее всего, следствием задержки при наборе текста, и поэтому не отражают особенности клавиатурного почерка. В результате возникла задача фильтрации потока длительностей на предмет выявления этих задержек. Поскольку указанная проблема упоминается в доступных источниках по тематике клавиатурного почерка, в дальнейшем предполагается подробнее проанализировать предлагаемые варианты ее решения на предмет практической применимости. Выяснено, что распределения длительностей межсимвольных интервалов не являются постоянными и могут изменяться как в зависимости от обстоятельств, при которых набирался текст, так и с течением времени. Отсюда сделан вывод о необходимости непрерывного учета этих изменений системой аутентификации. Таким образом, данные о характеристиках клавиатурного почерка должны постоянно обновляться, в противном случае система не будет обеспечивать надежного распознавания законного пользователя.

В результате проделанной работы разработан план исследований, на первом этапе которых необходимо выявить наиболее значимые признаки клавиатурного почерка, проявляющиеся при вводе пользователем текста, а также изучить их долговременную устойчивость. На этапе сбора исходных данных следует произвести классификацию пользователей по типу текстов, с которыми им приходится работать. Прежде всего следует разделить пользователей, которые вводят исключительно текст на русском языке, пользователей-программистов и пользователей, работающих с оконными формами СУБД в диалоговом режиме. Для количественной оценки устойчивости характеристик клавиатурного почерка предполагается использовать математические методы обнаружения разладки, а также методы последовательной проверки статистических гипотез. Эти же методы в дальнейшем предполагается использовать для обнаружения факта изменения клавиатурного почерка.